

AMENDMENTS TO THE CLAIMS:

This listing of claims will replace all prior versions, and listings of claims in the application:

1. (Currently Amended) An automated identification methodology for identification of table of content links in a given hyperdocument for assembling a document representation by gathering the content of hyperlinked pages pointed to by the identified table of contents comprising:

searching page data to create a list of links in the given hyperdocument;
analyzing each link in conjunction with each other link in the list of links to identify link pairings;

assembling link pairings in order to form clusters of links;
examining the links in the cluster of links for locality;

weeding out the links from the cluster of links which have properties that are not characteristic of intra-document links, to provide a resultant table of content set of identified candidate document pages; and,

grouping the content found in the resultant table of content set of candidate document pages by an automated system into a document representation stored in memory by the automated system; and,
for subsequent viewing or printing, or viewing on a display by a user, of the given hyperdocument/document representation.

2. (Original) The method of claim 1 wherein the step for analyzing each link further comprises determining a score for each link pairing.

3. (Original) The method of claim 2 wherein the scoring is determined by a proximity criteria.

4. (Original) The method of claim 2 wherein the scoring is determined by a similarity criteria.

5. (Original) The method of claim 2 wherein the scoring is determined by a regularity criteria.

6. (Currently Amended) A system identification methodology for assembling a document representation for subsequent viewing or printing of a given hyperlinked hyperdocument by gathering related hyperlinked page content comprising:

- performing a page-level link analysis that identifies those hyperlinks on a page linking to a candidate document page further comprising a methodology of:

- analyzing each link in conjunction with each other link to identify link pairings;

- assembling link pairings in order to form clusters of links; and,
 - examining the links in the cluster of links for locality;

- performing a recursive application of the page-level link analysis to the linked candidate document page and any further nested candidate document pages thereby identified, until a collective table of content set of identified candidate document pages is assembled;

- performing a document-level analysis that examines the collective table of content set of identified candidate document pages for grouping into one or more documents;

- examining the collective table of content set of identified candidate document pages to weed out links from the collective table of content set which have properties that are not characteristic of intra-document links, to provide a resultant set of identified candidate document pages; and,

- grouping the content found in the resultant set of candidate document pages by an automated system into a document representation stored in memory by the automated system; and,
for subsequent viewing or printing, or viewing on a display by a user, of the given hyperlinked hyperdocument document representation.

7. (Original) The method of claim 6 wherein the step for analyzing each link further comprises determining a score for each link pairing.

8. (Original) The method of claim 7 wherein the scoring is determined by a proximity criteria.

9. (Original) The method of claim 7 wherein the scoring is determined by a similarity criteria.

10. (Original) The method of claim 7 wherein the scoring is determined by a regularity criteria.

11. (Currently Amended) A system identification methodology for assembling a document representation for subsequent viewing or printing of a given hyperlinked hyperdocument by gathering related hyperlinked page content comprising:

performing a page-level link analysis that identifies those hyperlinks on a page linking to a candidate document page further comprising a methodology of:

searching page data to create a list of links in the hyperdocument;
analyzing each link in conjunction with each other link in the list of links to identify link pairings;
assembling link pairings in order to form clusters of links; and,
examining the links in the cluster of links for locality;

performing a recursive application of the page-level link analysis to the linked candidate document page and any further nested candidate document pages thereby identified, until a collective table of content set of identified candidate document pages is assembled; and,

performing a document-level analysis that examines the collective table of content set of identified candidate document pages for grouping into one or more documents

examining the collective table of content set of identified candidate document pages to weed out links from the collective table of content set which have properties that are not characteristic of intra-document links, to provide a resultant set of identified candidate document pages; and,

grouping the content found in the resultant set of candidate document pages by an automated system into a document representation stored in memory by the automated system; and,

~~for subsequent viewing or printing, or viewing on a display by a user, of the
given hyperlinked hyperdocument/document representation.~~

12. (Original) The method of claim 11 wherein the step for analyzing each link further comprises determining a score for each link pairing.
13. (Original) The method of claim 12 wherein the scoring is determined by a proximity criteria.
14. (Original) The method of claim 12 wherein the scoring is determined by a similarity criteria.
15. (Original) The method of claim 12 wherein the scoring is determined by a regularity criteria.